

REMARKS

Claims 1-32 are pending, of which, claims 1, 9, 13, 21, 26 and 29 are independent.

Claims 6, 8, 18 and 20 are amended.

Claim Objections

Claims 6, 8, 18 and 20 are objected to for including the term “normalizing.” (Office action, p. 2.) The Office action, however, provides guidance for overcoming the objection by pointing out the instances of occurrence of the term in the specification and cites to the specification’s description of the term. (Office action, pp. 2, 3.)

Applicants traverse the objection by submitting that the claims are clear in light of the specification. However, to expedite prosecution, the claims are being amended as suggested in the Office action.

Claim Rejections under 35 U.S.C. §103

Claims 1-4, 9-16, 21-24 and 26-31 are rejected under 35 U.S.C. §103(a) as being allegedly unpatentable over Hwang (U.S. Patent Application Publication No. 20020078090) in view of Foltz (U.S. Patent No. 6,356,864.)

Claims 5-8, 17-20, 25 and 32 are rejected under 35 U.S.C. §103(a) as being allegedly unpatentable over Hwang in view of Foltz and further in view of Boguraev (U.S. Patent No. 6,865,572.)

Applicants respectfully traverse these rejections in view of the following arguments.

Claim 1

Claim 1 recites “A method of creating a generic text summary of a document ... comprising ... creating a weighted document term-frequency vector for said document ... creating a weighted sentence term-frequency vector; computing a score for each said weighted

sentence term-frequency vector in accordance with relevance to said weighted document term-frequency vector; selecting a sentence for inclusion in said generic text summary in accordance with said computing ... representing high degree of relevance ... deleting said selected sentence from said document ... and generating the generic text summary based on the selected sentence." (Emphasis added.)

The Office action attempts to combine Hwang and Foltz to show the elements of claim 1. However, not all elements of claim 1 are taught or suggested by the two references.

For example, the Office action cites to paragraphs [0040] and [0041] of Hwang for teaching "deleting said selected sentence from said document" of claim 1. (Office action, p. 4.) As the Office action also states, in Hwang "The sentences determined for inclusion are then extracted ... along with any desired context information ... and merged." (Hwang, paragraph [0041].) There is no indication in Hwang that the "extracted" sentences are deleted. Further, in the method of Hwang, there is no reason or basis for conducting the extra step of "deleting" the extracted sentence because the relevance of the sentences is determined all at the same time and not in a progressive one by one fashion: "Initially, sentences to extract for summarization are selected (block 302). In one embodiment, all sentences in the original document that contain concept terms that would interest the user ... are marked for selection." (Hwang, paragraph [0036]; emphasis added.)

Further, conducting "creating a weighted document term-frequency vector ... creating a weighted sentence term-frequency vector ... [and] computing a score for each said weighted sentence term-frequency vector in accordance with relevance to said weighted document term-frequency vector" of claim 1 is novel and nonobvious in view of the cited references.

The Office action is citing to column 2, lines 20-25 and column 10, lines 15-25, of Foltz for the alleged teaching the above elements. (See, Office action, top p. 5 and top p. 6.) These passages teach that “A segment vector can be ... an entire reference ... [or] at least one sentence of a text” (Foltz, col. 10, lines 18-21.) Therefore, Foltz may disclose creating a “document ... vector” or a “sentence ... vector.” However, in broad terms, Foltz is a method for grading essays by comparing them to a standard text. As such, in general, Foltz is not interested in a document’s relationship with itself. And, it does not compare a part of a document to the whole of the same document. Unlike claim 1, if Foltz were to create a “document ... vector” it would next create another “document ... vector” and would not go to “creating a ... sentence ... vector.” Therefore, even if Foltz is interpreted as disclosing computing one of the “document ... vector” or a “sentence ... vector,” Foltz still does not teach or suggest computing both the “document ... vector” and “sentence ... vector” and computing a score for each “sentence ... vector” in accordance with relevance to the “... document vector,” as recited in claim 1.

Foltz, however, does have a provision for determining “coherence” of an essay that deals intra-document: “To determine the coherence of the student’s essay, the essay is divided into portions and a pseudo-object is computed for each portion of the essay. The cosine between the vector representing a first portion and the vector representing the following portion is a measure of the coherence between portions.” (Foltz, col. 4, lines 7-12; emphasis added.) Even this passage, however, compares a part to a subsequent and consecutive part. It does not compare a part to the whole. Here, Foltz would create a “sentence ... vector” followed by another “sentence ... vector” for a following sentence.

Accordingly, Hwang and Foltz, alone or combined do not teach or suggest “creating a weighted document term-frequency vector ... creating a weighted sentence term-frequency

vector ... computing a score for each said weighted sentence term-frequency vector in accordance with relevance to said weighted document term-frequency vector” of claim 1. This provides an additional reason of patentability of claim 1 over Hwang and Foltz. Thus, for all the foregoing reasons, claim 1 is patentable over Hwang, Foltz or any combination thereof.

Claim 9

Claim 9 recites “a vector generator for creating a weighted document term-frequency vector for said document and creating a weighted sentence term-frequency vector for each sentence in said document; a scoring engine for computing a score for each said weighted sentence term-frequency vector in accordance with relevance to said weighted document term-frequency vector; a selector for selecting a sentence for inclusion in said generic text summary in accordance with output results from said scoring engine; a document editor for deleting said selected sentence from said document and for eliminating terms in said selected sentence from said document.” (Emphasis added.)

Applicants submit that for reasons stated above with respect to claim 1, Hwang, alone or combined with Foltz, does not teach or suggest all elements of claim 9. Accordingly, this claim is believed to be patentable in view of the cited references.

Claim 13

Claim 13 recites “decomposing said document into individual sentences; forming a candidate sentence set from said individual sentences; for each of said individual sentences in said candidate sentence set, creating a weighted sentence term-frequency vector; creating a weighted document term-frequency vector for said document; for each of said individual sentences in said candidate sentence set, computing a relevance score for said weighted sentence term-frequency vector relative to said weighted document term-frequency vector; selecting a

sentence for inclusion in said generic text summary in accordance with said computing, wherein the selected sentence has the computed relevance score representing a high degree of relevance of the corresponding weighted sentence term-frequency vector to said weighted document term-frequency vector; deleting said selected sentence from said candidate sentence set." (Emphasis added.)

Applicants submit that for reasons stated above with respect to claim 1, Hwang, alone or combined with Foltz, does not teach or suggest all elements of claim 13. Accordingly, this claim is believed to be patentable in view of the cited references.

Claim 21

Claim 21 recites "A method of creating a generic text summary of a document; said method comprising: obtaining the document; constructing a terms-by-sentences matrix for said document; performing singular value decomposition on said terms-by-sentences matrix to obtain a singular value matrix and a right singular vector matrix, wherein each sentence in said document is represented by a column vector of a transpose of said right singular vector matrix; ranking each right singular vector in said right singular vector matrix; selecting a sentence for inclusion in said generic text summary in accordance with said ranking; and generating the generic text summary based on the selected sentence." (Emphasis added.)

The Office action is citing to Foltz, figure 1, column 6, lines 25-30 and column 7, lines 55 to column 9, line 40, for disclosing the elements of claim 21. (Office action, p. 9.)

Applicants submit that Foltz is merely reiterating M. Berry et al., Using Linear Algebra for Intelligent Information Retrieval, Tech. Rep. UT-CS-94-270, University of Tennessee, Computer Science Department, Dec. 1994, that is mentioned in paragraph 52 of the U.S. Patent Application Publication No. 2002/0138528, the published version of the specification of the

current Application. Singular Value Decomposition (SVD) is a standard linear algebra technique for finding a smaller-ranked, namely smaller, matrix, that is still a good approximation of a larger matrix. Berry uses SVD for analysis of texts and documents. Foltz uses Berry's idea and applies it to comparing a document to a standard document. Some aspects of the present invention use Berry's idea and apply it to text summarization. These aspects of the invention are at least as novel and nonobvious over Berry as Foltz was found to be by the USPTO.

First, Applicants submit that claim 21 uses SVD after "constructing a terms-by-sentences matrix for said document; performing singular value decomposition on said terms-by-sentences matrix." This is different from Foltz that, like Berry, is not interested in individual sentences, operates on entire documents, and finds the frequency of terms within the documents.

The above difference aside, Foltz does not teach or suggest "ranking each right singular vector in said right singular vector matrix; selecting a sentence for inclusion in said generic text summary in accordance with said ranking" of claim 21. Applicants submit that what is obtained by SVD has been known for some time. Application of SVD to document retrieval was presented by Berry. One use of what was discussed by Berry was patented by Foltz. Claim 21 is setting forth another different and distinguishable use that neither Berry nor Foltz appeared to have considered.

Hwang is merely cited because it pertains to document summarization and does not cure the deficiency of Foltz. Moreover, Hwang as stated by the Office action generates "a text summary ... focused on the user's interests." (Office action, p. 10; emphasis added.) Claim 21 recites "creating a generic text summary of a document." (emphasis added.)

As such, Applicants submit that Hwang, alone or combined with Foltz, does not teach or suggest all elements of claim 21. Accordingly, this claim is believed to be patentable in view of the cited references.

Claim 26

Claim 26 recites “a matrix generator for creating a terms-by-sentences matrix for said document; an SVD performer for performing singular value decomposition on said terms-by-sentences matrix to generate a singular value matrix and a right singular vector matrix; a vector analyzer for ranking each sentence in said terms-by-sentences matrix in accordance with index values with said right singular vector matrix; and a selector for selecting a sentence for inclusion in said generic text summary in accordance with output results from said vector analyzer; and a generic summary generator for generating the generic text summary based on the selected sentence.” (Emphasis added.)

Applicants submit that for reasons stated above with respect to claims 1 and 21, Hwang, alone or combined with Foltz, does not teach or suggest all elements of claim 26. Accordingly, this claim is believed to be patentable in view of the cited references.

Claim 29

Claim 29 recites “decomposing said document into individual sentences; forming a candidate sentence set from said individual sentences; constructing a terms-by-sentences matrix for said document; performing singular value decomposition on said terms-by-sentences matrix to obtain a singular value matrix and a right singular vector matrix, wherein each sentence in said candidate sentence set is represented by a column vector of a transpose of said right singular vector matrix; identifying a right singular vector from said right singular vector matrix; selecting

a sentence for inclusion in said generic text summary in accordance with said identifying;
selectively repeating said identifying and said selecting.” (Emphasis added.)

Applicants submit that for reasons stated above with respect to claims 1 and 21, Hwang, alone or combined with Foltz, does not teach or suggest all elements of claim 29. Accordingly, this claim is believed to be patentable in view of the cited references.

Independent Claims 1, 9, 13, 21, 26 and 29 in View of Boguraev

Boguraev was cited by the Office action for teaching the “weighting functions” used in some of the dependent claims. (Office action, p. 19.) However, the cited portions of this reference do not appear to cure the deficiencies of Hwang and Foltz in teaching or suggesting the elements of the independent claims. Independent claims 1, 9, 13, 21, 26 and 29, therefore, remain patentable over all cited references Hwang, Foltz and Boguraev, whether alone or in combination.

Dependent Claims

With respect to the rejection of dependent claims 2-8, 10-12, 14-20, 22-25, 27-28 and 30-32, while continuing to traverse the Examiner’s characterization of the teachings of the references used by the Examiner in rejecting these claims, Applicants respectfully submit that the rejection of these claims is rendered moot by the present amendments of the respective parent claims 1, 9, 13, 21, 26, and 29 and that these claims are patentable by definition, by virtue of their dependence upon their respective patentable independent claims.

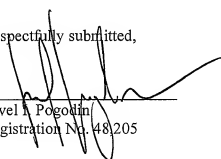
Conclusion

In view of the above, reconsideration and allowance of this application are now believed to be in order, and such actions are hereby solicited. If any points remain in issue which the

Examiner feels may be best resolved through a personal or telephone interview, the Examiner is kindly requested to contact the undersigned at the telephone number listed below.

The USPTO is directed and authorized to charge all required fees, except for the Issue Fee and the Publication Fee, to Deposit Account No. 19-4880. Please also credit any overpayments to said Deposit Account.

Respectfully submitted,



Pavel A. Pogodin
Registration No. 48,205

SUGHRUE MION, PLLC
Telephone: (202) 293-7060
Facsimile: (202) 293-7860

WASHINGTON DC OFFICE

23373

CUSTOMER NUMBER

Date: June 4, 2008